# Exposing the Science in Citizen Science: Fitness to Purpose and Intentional Design

Julia K. Parrish[1], Hillary Burgess[1], Jake F. Weltzin[2], Lucy Fortson[3], Andrea Wiggins[4], and Brooke Simmons[5]

[1]School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195, USA
[2]U.S. Geological Survey, 12201 Sunrise Valley Drive, Reston, VA 20192, USA
[3]?School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
[4]University of Nebraska at Omaha, 6001 Dodge Street, Omaha, NE 68182, USA
[5]University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

July 3, 2018. Accepted version

### Abstract

Citizen science is a growing phenomenon. With millions of people involved and billions of in-kind dollars contributed annually, this broad extent, fine grain approach to data collection should be garnering enthusiastic support in the mainstream science and higher education communities. However, many academic researchers demonstrate distinct biases against the use of citizen science as a source of rigorous information. To engage the public in scientific research, and the research community in the practice of citizen science, a mutual understanding is needed of accepted quality standards in science, and the corresponding specifics of project design and implementation when working with a broad public base. We define a science-based typology focused on the degree to which projects deliver the type(s) and quality of data/work needed to produce valid scientific outcomes directly useful in science and natural resource management. Where project intent includes direct contribution to science and the public is actively involved either virtually or hands-on, we examine the measures of quality assurance (methods to increase data quality during the design and implementation phases of a project) and quality control (post hoc methods to increase the quality of scientific outcomes). We suggest that high quality science can be produced with massive, largely one–off, participation if data collection is simple and quality control includes algorithm voting, statistical pruning, and/or computational modeling. Small to mid-scale projects engaging participants in repeated, often complex, sampling can advance quality through expert-led training and well-designed materials, and through independent verification. Both approaches—simplification at scale and complexity with care—generate more robust science outcomes.

## Introduction

On December 22, 2014, Virginia started her sixth beached bird survey near Ocean Shores, Washington. Trained only 2 months previously, she was still on the learning curve. In fact, she got a lot of practice that day. Virginia and her survey partner found 425 carcasses in less than a kilometer, and photographed, tagged, and identified all of them. This single survey marked the peak of the largest marine bird mass mortality event ever documented in the Pacific Northwest of the United States (Jones et al. 2018). A documentation only possible because more than 500 trained participants of the BeachCOMBERS, BeachWatch, and COASST beached bird survey programs conducted over 1650 standardized, effort-controlled surveys at 264 sites from Morro Bay, CA to Neah Bay, WA. At the same time, program experts verified carcass identification from the collected evidence (photographs, standard measurements, foot type). Finally, almost 20 scientists, including oceanographers, atmospheric scientists, marine ecologists, veterinary pathologists, and seabird biologists brought their expertise to bear in determining the extent, intensity, and causality of the event. In this story, citizen science and science are synonymous. Is this the norm, or the exception? In this paper, we examine the attributes of citizen science leading to rigorous and robust science.

We define citizen science as projects in which members of the public engage directly in research developed by or with scientists to address particular questions and/or issues (Irwin 1995; Bonney et al. 2009a). Because the term "citizen" can be politically problematic and the term "volunteer" is not always appropriate, we refer to individuals directly involved in citizen science projects and not including project staff as "participants." Within natural science, fields utilizing citizen science already include: archaeology (Bovy et al. 2016), astronomy (Fortson et al. 2012), biochemistry (Eiben et al. 2012), ecology (Dickinson et al. 2010), geography (Goodchild 2007), geology (Powell et al. 2013), and oceanography (Hays et al. 2005). This diversity might suggest that academic and professional science is broadly accepting of public involvement; however, recent studies indicate that the mainstream scientific community remains skeptical of the public as a trusted source of scientific information (Riesch and Potter 2014; Burgess et al. 2017). In many cases, these misgivings are rooted in the demonstration that non-experts in a citizen science program do not always perform a scientific task (usually data collection) to the standards desired by researchers. Thus, the evidence that some citizen science programs produce high quality data of immediate use to science (e.g., Cooper et al. 2014; Swanson et al. 2016) does not translate into the conclusion that all citizen science programs can.

## Defining the goals

Many citizen science projects assert production of data in service to science or resource management as a goal. Theobald et al. (2015) found that 97% of 388 surveyed biodiversity citizen science projects stated their primary goal was to contribute to science and/or advance scientific understanding. However, only 12% of projects had demonstrably contributed to a science-focused peer-reviewed publication (one measure of scientific contribution). Even if this publication rate is underreported due to "cryptic" use of the term citizen science only outside of the abstract and keywords if at all (Cooper et al. 2014), the discrepancy suggests that there may be large differences in what project managers, and research scientists, consider evidence of scientific use. In assessing the potential for bonafide science as an outcome of citizen science, we invoke the concept of fitness to use or fitness to purpose (Juran 1951), or the degree to which the quality-related elements or activities of an organization—here a citizen science project—can result in the declared purpose. Simply put, projects claiming science as a primary goal or "purpose" should adhere to accepted quality standards within science (Wiggins et al. 2018).

However, science is not the only goal of citizen science. Other common goals include education, community empowerment, and personal fulfillment. Science education and/or increasing science literacy has long been a major thrust of citizen science programming (Bonney et al. 2009b; Wiggins and Crowston 2011). Community goals, often related to environmental or social justice issues, are an explicit outcome of community-based, community-driven, and participant action research projects (Wilderman et al. 2004; Cooper et al. 2007; Danielsen et al. 2009). And for the individual participant, personal fulfillment can include learning goals, the desire to contribute to science, or simply engaging in something enjoyable or fun (Raddick et al. 2010; Wright et al. 2015).

While we recognize the value of citizen science to both personal and societal outcomes, this paper explores strategies for better ensuring projects can meet declared goals based on scientific outcomes (i.e., optimizing project fitness to scientific purpose). Here we distinguish between the practice of science (including authentic science experiences on the part of the participants) and science outcomes (new information or knowledge, or applied work based on a scientific understanding of how the world works), where the latter must include the former, but the reverse is not the case. Our goal is to facilitate both acceptance and use of citizen science by the professional science community, and intentional design of projects with science as a primary objective. To that end, we: (1) present a science-focused typology that differentiates projects based on intent and activity; (2) define a process workflow to help identify design nexus points for science-focused projects; and (3) discuss quality control approaches to maximize data quality as a function of project scale and complexity.

## A science-based typology of citizen science

Existing typologies of citizen science pivot on the degree to which participants are involved in tasks other than data collection. Bonney et al. (2009b) posited three points of project design along a continuum of interaction between scientist and participant. Contributory projects—also referred to as virtual and/or investigative projects (Wiggins and Crowston 2011), externally-driven monitoring with local data collectors (Danielsen et al. 2009), or distributed intelligence (Haklay 2013)—are designed by the mainstream science community with the role of data producer assigned to the public. At the other end of the continuum are co-created projects which involve participants in all stages of the scientific process, and are often associated with particular communities and specific concerns such as air or water quality, as in "extreme citizen science" or community-based participatory science focused on highly marginalized and often remote populations (Haklay 2013; Stevens et al. 2014). In fact, there are a range of projects which confer increasing power and project ownership to non-scientist participants including autonomous local monitoring (Danielsen et al. 2009), community-based participatory research (Wilderman et al. 2004), and more generally "action" projects (Wiggins and Crowston 2011). What often sets these projects apart is the explicit movement of project results into the sphere of decision-making and governance. In between these poles are collaborative projects expanding participants roles beyond data collector, from contributing to iterative versions of data collection protocols and training of new recruits, to results interpretation and defining the next phase(s) of the research (Cooper et al. 2007).

What is apparent about most of these typologies is that they are centered on the roles and degree of control accorded to professional scientists versus other participants. We suggest that a science-focused typology aimed at classifying projects according to their potential for inclusion in scientific research and science-based decision-making is also needed to guide the scientific community in identifying projects applicable to their work. In lieu of a meta-analysis systematically reviewing attributes of all citizen science projects (e.g., the 1800 projects currently listed in SciStarter.org), we generated our schema through an iterative process that extended a framework presented at the Waypoints of Science: Scaling Design, Development and Delivery of Citizen Science for Bonafide Science symposium held at the Citizen Science Association meeting in 2017. Iterations were tested against: (1) all projects (unique projects 1/4 80) highlighted as examples in all previous literature proffering a typology or categorization of citizen science (i.e., see references above), (2) the 388 biodiversity citizen science projects collected in the Theobald et al. (2015) meta-analysis, (3) projects managed directly by the authors, and projects associated with and/or analogous to or duplicative of those projects (e.g., all projects focused on beach habitats; projects focused on documenting phenology), and (4) all projects on data collection platforms managed by the authors (e.g., in the Zooniverse). In total, over 500 projects were tested against our typology.

The right-hand branch of Fig. 1—no/minimal data—is defined by projects for which the primary intent is not data collection at a level or scale needed to address an issue or question of scientific interest. Education and awareness projects may well bring members of the public into direct contact with practicing scientists for the first time, and may provide individuals with authentic scientific experiences, without contributing to the advancement of science. Examples include the Lost Ladybug Project (Gardiner et al. 2012) which focuses on youth programs to identify native and invasive ladybugs, and the youth-focused intertidal project Long-term Monitoring Program and Experiential Training for Students, or LiMPETS (Ballard et al. 2017). In both of these examples, hundreds of middle school students annually gain authentic science experiences, become more aware of scientific practice and environmental issues, and may gain agency (permission to act) and expand their identity through participation (Ballard et al. 2017). However, standardized, effort-controlled, verifiable data at a spatio-temporal scale equivalent to questions of scientific interest are rarely produced. Non-data collection tasks include a broad swath of activities where participants may be deeply engaged in assistance toward a goal that does not require the collection or processing of information, as in conservation action and restoration projects (Bruyere and Rappe 2007).

The left-hand branch—data generated—separates out projects where the primary intent is the creation of information, or data in service of a scientific goal. We define "data" as an abstraction—a measurement, classification, and/or count that individually or collectively characterizes an object, phenomenon, or state—as well as the thing itself, as in a sample. First, we divide projects by whether the participant is directly engaged in thinking, or is giving tacit permission for the use of "information and communication technologies" (Wiggins and Crowston 2011). Passive participation ranges from computation, or the use of networked desktops and laptops to parallel process discrete "work packages" as part of big data projects (e.g., SETI@home, Rosetta@home, climate-prediction.net), to sensing, defined as personally carrying and/or housing automated sensors which report data directly (e.g., Quake-Catcher Network, where participants host seismic sensors on their laptops; Cochran et al. 2009). Although science is clearly being accomplished in both cases, the participant is passive in the sense of a non-thinking contribution which can be accomplished without specific understanding of how their participation contributes to science.

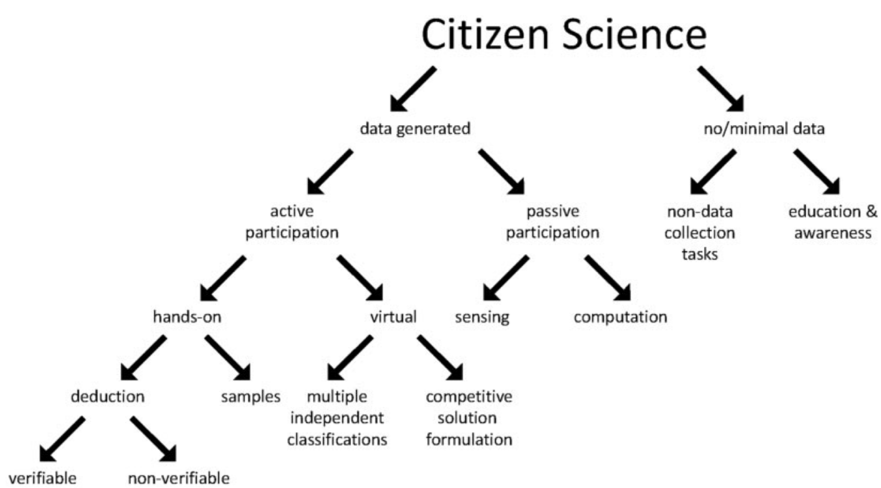


Figure 1: A typology of citizen science separating projects according to scientific intent and participant activities. The first branching separates projects by primary goal: generation of science outcomes (e.g., data generated) or other goals (e.g., education, community empowerment, or personal fulfillment) for which data generation is possible but not necessary. Note any single branch point does not define mutually exclusive space (e.g., awareness/education is possible on the left side of the tree, and deduction can be accomplished virtually).

By contrast, active participation requires participants to engage directly in one or more of the tasks of the scientific process. Types of activity can be divided into physical hands-on work and virtual citizen science—where the latter is conducted entirely through a computer interface, often online, whether that is situated in a kiosk at a visitor's center or in a science museum, at home, or on a mobile device. Virtual citizen science capitalizes on crowdsourcing, a distributed production model that makes an open call for contributions from a large, undefined network of people (Howe 2006) to achieve both faster task accomplishment and higher project-wide accuracy with no precondition or expectation of long-term engagement.

Two basic approaches to crowdsourcing in service of science include: multiple independent classifications and competitive solution formulation. In the former, the accuracy of the individual participant is secondary to the "wisdom of the crowd" emerging through the use of voting or aggregation algorithms (Fortson et al. 2012). Advanced algorithms account for individual performance, assigning additional weight to responses from participants who are more accurate, and/or who contribute more (e.g., Marshall et al. 2016). Zooniverse—an online, crowd-source classification platform currently hosting ˜80 projects is the exemplar. Zooniverse participants can choose to classify everything from camera-trapped mammals in East Africa (Snapshot Serengeti) to feather color from digital stills of museum specimens (Project Plumage) to leaves on growing plants (Leaf Targeting). By contrast, competitive solution formulation uses the crowd to find the single best participant, as in the protein structure game Foldit (Khatib et al. 2011) or the multiple sequence alignment game Phylo (Kawrykow et al. 2012). Task performance is tied directly to recognition and thus a degree of competition (e.g., Greenhill et al 2014), and the "game" may become relatively distinct from the underlying science.

Finally, hands-on citizen science is typified by a wide range of projects from laboratory-based work to field-based environmental science. These projects include both monitoring and experimental studies, all of which require physical collection of data. Sample collection includes direct contact with the sampled material, as in SoundCitizen, a water quality project requiring participants to send in water samples for laboratory analysis (Keil et al. 2011); and/or may simply be a geo-referenced, time-stamped photograph, as in CrowdWater, which collects hydrological data based on photographs (Seibert et al. 2017). In deduction, a decision is made based on the original data or evidence collected (e.g., species identification based on morphological characters), as in the fish identification dive program Reef Environmental Education Foundation Fish Survey Project (REEF; Thorson et al. 2014). For verifiable deductions, the decision reached by a participant can be independently verified; that is, an expert can evaluate the collected evidence, as is the case with Earthwatch, where experts are on-site with participants (Chandler et al. 2012). Non-verifiable deductions can still have high scientific value, especially when the volume or scale of data collected is high or large, as is the case for the Christmas Bird Count, or eBird (Sullivan et al. 2014). In virtual projects, verification solutions are implicit in the crowdsource approach.

## Designing for science and citizen science

An increasing body of literature suggests that non-professional participants engaged in hands-on, deductive citizen science may underperform relative to professionals. For example, project participants tend to under-report common species and over-report rare species (Kremen et al. 2011; Paul et al. 2014). Participants over-report easy-to-identify, flashy, brightly colored or especially charismatic species (Ward 2014; Boakes et al. 2016) and under-report cryptic species (Cox et al. 2015). Non-professionals are less likely to master non-visual survey methods (e.g., acoustic surveys, scat surveys; Moyer-Horner et al. 2012), and are more likely to collect information non-systematically across the landscape (Boakes et al. 2016). In contrast, a meta-analysis of 509 ecological and environmental citizen science projects (Pocock et al. 2017) found that "best quality of data" was associated with in-person training, production of associated materials (e.g., a protocol), and the use of specialized equipment for data collection.

For citizen science to become an accepted form of bona-fide science, intentional design with attention to data quality is essential, including measures of quality assurance (the procedures to enhance data quality undertaken before and during data collection) and methods of quality control (the processes for improving quality after data collection). Burgess et al. (2017) found that biodiversity scientists overwhelmingly agreed on the following quality assurance measures for field-collected data: documentation of sampling location, time, and date; effort control via known area and/or time envelope of sampling; verifiable data; and data collection personnel trained by an expert.

We abstracted the scientific process as a series of steps (left side, Fig. 2) from project design through to publication and use, that can be understood as necessary in both science (flowchart in gray, Fig. 2) and citizen science (flowchart in white, Fig. 2). The design of any scientific project design involves the selection of a sampling scale and a level of precision for data collection that match the question or issue at hand, as well as selecting a minimum sample size (N floor) that addresses the variability inherent in the system. Once the data are collected, the post-processing step involves refining an analytic approach suited to the data and the question. The final step in science is presenting the work in a peer-reviewed publication.

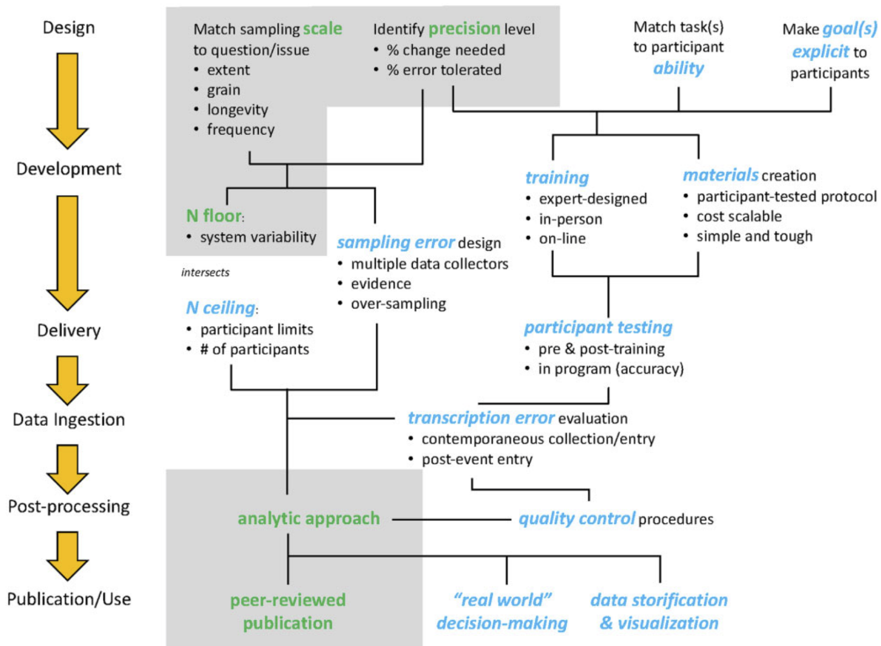


Figure 2: The steps of science (listed sequentially at left) outlined as a flowchart. At each stage, the necessary elements inherent in all science projects are highlighted in bold print and encased within the gray box across stages. Additional elements specific to citizen science are highlighted in bold italics, and fall outside the gray box.

## Quality assurance in citizen science

Citizen science as a method of science is not different, but requires additional attention to aspects of quality assurance. During project design, intentional recruitment of target audiences can be key to success. Individuals are differentially attracted to projects based on personal values and shared goals (Evans et al. 2005; Rotman et al. 2012). Thus, making project goals explicit allows individuals who may have different, even antithetical, goals to consider whether their needs are being met, perhaps selecting another project more closely aligned to their own world-view. Attention to ability, or level of content knowledge and skill development as novice participants, is also essential. Projects are variably accessible relative to physical ability, economic status, and time available among other features (Pandya 2012). Whether recruits can accomplish the work will also vary as a function of their "distance" from the content and the complexity of the tasks (Jung et al. 2005; Kosmala et al. 2016). For instance, while some projects attract hobbyists with a high degree of skill and little need of formal training (e.g., birders, amateur astronomers—Jones et al. 2017a), many projects attract a broad swath of interested non-experts with little-to-no a priori training (Kelling et al. 2015).

Within the realm of citizen science, project development follows from the intersection of participant ability and the sampling precision required by the project, and includes two types of interaction with participants: training and participant-specific materials. While scientists prefer citizen science data collected by projects with in-person expert training (Burgess et al. 2017), online trainings can also be effective (Masters et al. 2016), and may be the only way to scale projects beyond local-to-regional geographies. Project materials include, at a minimum, a well-developed protocol outlining all of the steps needed to perform tasks successfully, and project-specific tools (e.g., measuring equipment, data sheets). Parrish et al. (2017) suggest serial refinement of project materials—(in this case, a field key to beached birds) by non-professional, non-experts in the target audience in collaboration with project scientists—to identify and replace or explain jargon and otherwise clarify materials. Co-creation and/or refinement of the training, protocol, and associated data collection materials among scientists, project staff, and project participants can improve both data quality and participant retention (Kim et al. 2011). Attention to cost-effectiveness, including both the price of provided materials and their durability, is important because the scaled success of a project— recruiting thousands of participants—should not cause its financial failure nor exclude potential participants in disadvantaged circumstances.

In the delivery of the project, quality assurance can be affected through participant testing and attention to sampling. Testing participant knowledge can be used to ensure that trainings are successful in delivering both content and skill (e.g., pre-post testing surrounding a training), as well as to ensure continued quality as participants engage in the practice of project tasks; that is, do the work. For online image classification projects, inserting a certain proportion of images where the answer is already known can create an accuracy baseline for each participant. Such evaluation built directly into the normal flow

of activities (i.e., embedded assessment) can also support timely feedback. For participants, understanding what they are doing incorrectly and how to correct it, as well as recognition of correctly accomplished tasks, can be empowering and lead to increased retention (Haywood et al. 2016). For project designers, understanding process breakdowns is essential for adaptively managing project training and materials to maximize data quality, as well as to understand the types and levels of error resulting from hundreds to thousands of data collectors. Although minimum sample size is set by system variability, maximum sample size (N ceiling) should be set relative to what individual participants can reasonably be expected to contribute, multiplied by the number of participants (minimally) in the program. Because citizen science is, by definition, the work of the many, attending to the sampling error inherent in this design is important, and may further increase sampling needs depending on whether participants are collecting deductive data that is (or isn't) backed up by evidence.

Data ingestion is automatic in some projects (e.g., all passive participation and some virtual, and sample collection projects) such that transcription error is non-existent. Virtual projects focused on classification (e.g., projects within the Zooniverse) minimize transcription error via the crowdsource design of multiple, independent classifiers for each task. However, hands-on projects may provide participants with the opportunity to input data they collect, introducing another source of error in the data. Mobile technologies may offer solutions by automatically logging some data (e.g., date, time, location, limited environmental data, and photographic evidence).

## Quality control in citizen science

Within citizen science, post-processing prior to analysis offers many possibilities for post-hoc improvement to data via quality control procedures, even in cases where quality assurance has been relatively weak. In Fig. 3, we conceptualize citizen science projects from those featuring relatively simple tasks requiring little-to-no deductive reasoning on the part of the participant (e.g., collecting a water sample, collecting a photograph sample) to those requiring participants to engage in complicated work requiring advanced training, deductive reasoning, mastery through practice, and/or a mental model of the system (e.g., species identification, performing chemical analyses on water quality samples). Orthogonal to the axis of task complexity, we array projects as a function of scale, from local projects with relatively few participants to projects that span regions (e.g., large marine ecosystems, countries, or continents) up to— at least theoretically—the globe. While not completely interchangeable, projects with a larger geographic extent also tend to be those with higher participant numbers (Theobald et al. 2015). Virtual projects, which are effectively aspatial, can similarly scale in participant numbers and total tasks completed.

For simple tasks (left side of Fig. 3), data quality can be improved by "outsourcing" the thinking to scientists, that is, restricting citizen involvement to straightforward sample collection tasks while scientists receive, verify, catalog, and analyze the samples and the resulting data (i.e., do the thinking). In the case of virtual projects with numbers of participants (upper left quadrant of Fig. 3), data quality can be improved via crowdsourcing tasks to multiple individuals, with task completion automatically based on algorithm voting or consensus metrics (e.g., species identification projects on the Zooniverse platform). For example, Swanson et al. (2016) found that crowdsourced (>10 people classifying an image) identifications of images in Snapshot Serengeti were slightly (97.9%) more accurate than even expert identifications (96.6%). Algorithms can also identify individual players who are particularly adept, or inept, and assign coefficients accordingly (Hines et al. 2015), creating more accurate data (Marshall et al. 2016)—akin to statistical pruning. While outsourcing is constrained by scientific resource time to smaller projects, crowdsourcing supports very large projects with millions of images to be processed (e.g., Lintott et al. 2008). Here, even inaccurate answers can prove valuable information if a given participant's bias is systematic (Masters et al. 2016).

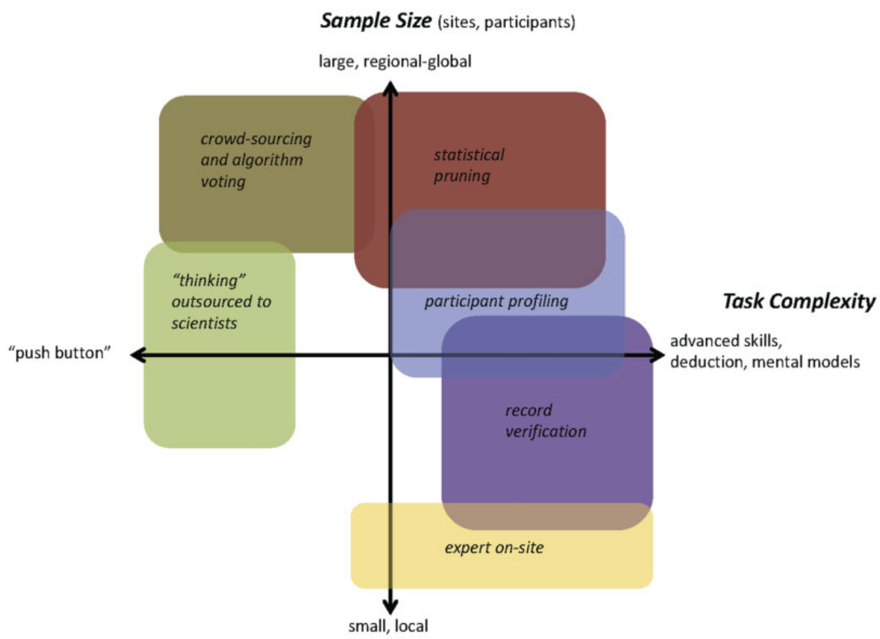


Figure 3: Approaches to quality control in citizen science as a function of the scale and complexity of the task(s) performed by participants. Shading is used to visually highlight the different approaches. Regions of overlap indicate intersections of task complexity and sample size within which multiple solutions might be found.

As task complexity increases at small project scales (lower right quadrant of Fig. 3), options for quality control shift toward expert intervention. On-site expert facilitation and mentoring is exemplified by Earthwatch where scientists train, mentor, and remain on-site with participants throughout the tenure of the project (Chandler et al. 2017). In independent record verification, participants' deductions are subsequently verified via photographs or specimens. For example, in the COASST program all species identifications (marine birds) are independently verified by experts via participant-collected primary evidence (foot type, standardized measurements, scaled dorsal, and ventral photographs), a process that improves identification to species level from 83% (participant rate) to 92% (Parrish et al. 2017). Verification can also proceed at the local phenomenological level, as in tracking the invasion front of the Asian tiger mosquito (Aedes albopictus) in Spain, where participant reports via the Mosquito Alert app were independently verified via ovitrapping (Palmer et al. 2017).

As project scales increase to continental and beyond (upper right quadrant of Fig. 3), quality control of individual data points may be less practical as volume prohibits comprehensive expert review, but statistical pruning, flagging, and other post hoc techniques can weed out anomalous data points (e.g., mixed effect models and machine learning: Bird et al. 2014; false-positive occupancy models: Pillay et al. 2014), or computational models can be used to create smoothed, interpolated versions of the original data (e.g., spatiotemporal exploratory models: Hochachka et al. 2012). In between, participant profiling (e.g., trust metrics: Hunter et al. 2013; occupancy-detection-experience model: Hochachka et al. 2012) can be used to winnow or weight data based on participants' known performance levels; however, this approach can introduce difficult decisions about the ethics of selective data use.

## Use beyond science

For most academics, the ultimate step is dissemination of results into the scientific literature (i.e., "publication/use" step in Fig. 2), simultaneously validating the work through review by scientific peers while daylighting the work to the larger scientific community. However, long-term maintenance of a citizen science project requires two additional and on-going steps: demonstrating that science is applied as promised, and sharing the results with participants (Cox et al. 2015). For some projects, taking results directly into "real world" decision-making processes (e.g., conservation, resource management) is the social contract that contributors make as a precondition for participation (Haywood et al. 2016). For place-based, environmental justice projects, such decision-making is the primary, even exclusive, goal (Haklay 2013).

Finally, returning "results-at-scale" to participants in suitable text and graphical forms (i.e., data storification and visualization, Fig. 2) can be essential to participant retention (Cox et al. 2015). Species migration (e.g., eBird—Sullivan et al. 2014), the timing of spring flowering (e.g., National Phenology Network—Rosemartin et al. 2014), the occurrence and location of extreme weather (e.g., CoCoRaHS—Gochis et al. 2015), the spread of disease across a population (e.g., Sea Star Wasting— Montecino-Latorre et al. 2016), the extent of a marine bird mass mortality event (e.g., COASST—Jones et al. 2017b)—these "data stories" are all patterns that transcend the ability of a single participant to directly observe the emergent pattern. Without these stories, participants cannot "see" their own data as contributing to the greater whole, and may be unaware of actual data uses. With these stories, participants refer to the work as "purposeful and powerful" and may be energized to take action, from continued engagement to calling for conservation stewardship or other resource management outcomes (Haywood et al. 2016).

## Conclusions

Citizen science progresses through the actions of the many. The collective work of hundreds to hundreds-of-thousands creates datasets that bound phenomena and address issues of scientific and management interest at spatio-temporal scales otherwise unattainable (Theobald et al. 2015). With this promise comes responsibility:

- from the scientific community to erase or at least understand bias and to embrace well-designed, scale-, and content-appropriate projects as a valid source of information;
- from project designers to attend to the specifics of quality assurance and quality control needed to produce rigorous, high quality data, if science is the primary goal;
- from project owners and managers to honestly advertise the type of project, depth of participant engagement, and quality and limitations of project data, and to ensure fitness to declared purpose;

- from participating scientists to follow through on data use and data stories providing both the scientific community and the participant corps and their communities with results-at-scale;
- and from participants to choose projects wisely according to their values and goals, to contribute as much and as well as they can, and to hold project managers to their declarations of purpose or intent.

Without judgment, we suggest the use of a science-based typology to sort existing projects will increase the "honest signaling" needed to help the mainstream science community see and understand citizen science as a bonafide method for generating legitimate scientific outcomes. Furthermore, the degree to which the individual participant: (1) understands and values the precision and accuracy required of the task(s) they are performing; (2) applies "thinking" skills requiring mastery of simple tasks to successfully perform more complex ones (e.g., species identification); and (3) can literally "see" their work (data collection or otherwise) within the larger context defined by the science at scale, will structure their degree of engagement and will impact data quality. Because task performance is often dependent on accrued experience within a project (Kelling et al. 2015; Kosmala et al. 2016), the strategies we have outlined herein (i.e., Fig. 2) support the "learning curve" and improve retention by providing transparency about project goals and data quality processes that match fitness to purpose (Juran 1951). Pocock et al. (2017) found that the recent 10% per annum growth rate in ecological and environmental citizen science has primarily been realized through online projects with mass, often short-term, participation in low-complexity data collection. Growth in field-based, hands-on approaches is more difficult, but can return data on global change impacts from climate to disease to invasions and ecosystem change (Theobald et al. 2015). Thus, we argue that both approaches—simplification at scale, and complexity with care—are valid and valuable strategies for citizen science projects to generate rigorous and robust science outcomes.

# Acknowledgments

Contribution to the dedicated SICB volume on "Science in the Public Eye" January 2018 meeting symposium. Portions of this manuscript were conceived at a citizen science workshop sponsored by the US Department of Agriculture-Forest Service and Schoodic Institute. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US government.

# Funding

A.W. was supported in part by U.S. Geological Survey (USGS) Cooperative Agreement G16AC00267 and National Science Foundation (NSF) grant 0830944. J.K.P. and H.B. were partially supported by NSF EHR/DRL awards 1114734 and 1322820, and Washington Sea Grant award R/ OLWD-1. B.S. acknowledges support from the National Aeronautics and Space Administration (NASA) through Einstein Postdoctoral Fellowship Award Number PF5-160143 issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS8-03060. L.F. acknowledges partial support from NSF IIS 1619177.

# References

Ballard HL, Dixon CGH, Harris EM. 2017. Youth-focused citizen science: examining the role of environmental science learning and agency for conservation. Biol Conserv 208:65–75.

Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Stuart-Smith RD, Wotherspoon S, Krkosek M, Stuart-Smith JF. 2014. Statistical solutions for error and bias in global citizen science datasets. Biol Conserv 173:144–54.

Boakes EH, Gliozzo G, Seymour V, Harvey M, Smith C, Roy DB, Haklay M. 2016. Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. Sci Rep 6:33051.

Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips TB, Rosenberg KV, Shirk J. 2009a. Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59:977–84.

Bonney R, Ballard H, Jordan H, McCallie E, Phillips T, Shirk J, Wilderman C. 2009b. Public participation in scientific research: defining the field and assessing its potential for informal science education. A CAISE Inquiry Group Report. Washington, DC: Center for Advancement of Informal Science Education (CAISE).

Bovy KM, Watson JE, Dolliver J, Parrish JK. 2016. Distinguishing offshore bird hunting from beach scavenging in archaeological contexts: the value of modern beach surveys. J Archaeol Sci 70:35–47.

Bruyere B, Rappe S. 2007. Identifying the motivations of environmental volunteers. J Environ Plann Manage 50:503–16. Burgess H, DeBey LB, Froehlich H, Schmidt N, Theobald

EJ, Ettinger AK, HilleRisLambers J, Tewksbury J, Parrish JK. 2017. The science of citizen science: exploring barriers to use as a primary research tool. Biol Conserv 208:113–20.

Chandler M, Bebber DP, Castro S, Lowman MD, Muoria P, Oguge N, Rubenstein DI. 2012. International citizen science: making the local global. Front Ecol Environ 10:328–31.

Chandler M, Rullman S, Cousins J, Esmail N, Begin E, Venicx G, Eisenberg C, Studer M. 2017. Contributions to publications and management plans from 7 years of citizen science: use of a novel tool on Earthwatch-supported projects. Biol Conserv 208:163–73.

Cochran ES, Lawrence JF, Christensen C, Jakka RS. 2009. The quake-catcher network: citizen science expanding seismic horizons. Seismol Res Lett 80:26–30.

Cooper CB, Dickinson J, Phillips T, Bonney R. 2007. Citizen science as a tool for conservation in residential ecosystems. Ecol Soc 12:11–22.

Cooper CB, Shir J, Zuckerberg B. 2014. The invisible prevalence of citizen science in global research: migratory birds and climate change. PLoS One 9:e106508.

Cox J, Oh EY, Simmons B, Lintott C, Masters K, Graham G, Greenhill A, Holmes K. 2015. Defining and measuring success in online citizen science: a case study of zooniverse projects. Comput Sci Eng 17:28–41.

Danielsen F, Burgess N, Balmford A, Donald P, Funder M, Jones J, Alviola P, Balete D, Blomley T, Brashares J, et al. 2009. Local participation in natural resource monitoring: a characterization of approaches. Conserv Biol 23:31–42.

Dickinson JL, Zuckerberg B, Bonter DN. 2010. Citizen science as an ecological research tool: challenges and benefits. Annu Rev Ecol Evol Syst 41:149–72.

Eiben CB, Siegel JB, Bale JB, Cooper S, Khatib F, Shen BW, Players F, Stoddard BL, Popovic Z, Baker D. 2012. Increased diels-alderase activity through backbone remodeling guided by Foldit players. Nat Biotechnol 30:190–2.

Evans C, Abrams E, Reitsma R, Roux K, Salmonsen L, Marra PP. 2005. The neighborhood nestwatch program: participant outcomes of a citizen-science ecological research project. Conserv Biol 19:589–94.

Fortson LF, Masters K, Nichol R, Borne K, Edmondson EM, Lintott C, Raddick J, Schawinski K, Wallin J. 2012. Galaxy zoo: morphological classification and citizen science. In: Way MJ, Scargle JD, Ali KM, Srivastava AN, editors. Advances in machine learning and data mining for astronomy. Boca Raton (FL): CRC Press, Taylor & Francis Group. p. 213–36.

Gardiner MM, Allee LL, Brown PMJ, Losey JE, Roy HE, Smyth RR. 2012. Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. Front Ecol Environ 10:471–6.

Gochis D, Schumacher R, Friedrich K, Doesken N, Kelsch M, Sun J, Ikeda K, Lindsey D, Wood A, Dolan B. 2015. The great Colorado flood of September 2013. Bull Am Meterol Soc 96:1461–87.

Goodchild M. 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69:211–21.

Greenhill A, Holmes K, Lintott C, Simmons B, Masters K, Cox J, Graham G. 2014. Playing with science: gamised aspects of gamification found on the Online Citizen Science Project—Zooniverse. In: Dickinson P, editor. GAMEON 2014. Lincoln: EUROSIS—The European Multidisciplinary Society for Modelling and Simulation Technology. p. 15–24.

Haklay M. 2013. Citizen science and volunteered geographic information: overview and typology of participation. In: Sui DZ, Elwood S, Goodchild MF, editors. Crowdsourcing geographic knowledge: voluteered geographic information (VGI) in theory and practice. Berlin: Springer. p. 105–22. Hays GC, Richardson AJ, Robinson C. 2005. Climate change

and marine plankton. Trends Ecol Evol 20:337–44. Haywood B, Parrish JK, Dolliver J. 2016. Place-based, data-rich citizen science as a precursor for conservation action. Conserv Biol 30:476–86.

Hines G, Kosmala M, Swanson A, Lintott CJ. 2015. Aggregating user input in ecology citizen science projects. Proceedings of the Twenty-Seventh Conference on Innovative Applications of Artificial Intelligence. Palo Alto (CA): AAAI Press. p. 3975–80.

Hochachka WM, Fink D, Hutchinson RA, Sheldon D, Wong WK, Kelling S. 2012. Data-intensive science applied to broad-scale citizen science. Trends Ecol Evol 27:130–7.

Howe J. 2006. The rise of crowdsourcing. Wired Magazine, June 2006.

Hunter J, Alabri A, van Ingen C. 2013. Assessing the quality and trustworthiness of citizen science data. Concurr Comput 25:454–66.

Irwin A. 1995. Citizen science: a study of people, expertise, and sustainable development. New York (NY): Routledge. Jones MG, Corin EN, Andre T, Childers GM, Stevens V. 2017a. Factors contributing to lifelong science learning: amateur astronomers and birders. J Res Sci Teach 54:412–33.

Jones TJ, Parrish JK, Punt AE, Trainer VL, Kudela R, Lang J, Brancato MS, Odell A, Hickey B. 2017b. A mass mortality event of marine birds in the Northeast Pacific caused by Akashiwo sanguinea. Mar Ecol Progr Ser 579:111–27.

Jones TJ, Parrish JK, Peterson WT, Bjorkstedt EP, Bond NA, Ballance LT, Bowes V, Hipfner JM, Burgess HK, Dolliver JE, et al. 2018. Massive mortality of a planktivorous seabird in response to a marine heatwave. Geophys Res Lett 45:3193–202.

Jung W, Olfman L, Ryan T, Park YT. 2005. An experimental study of the effects of contextual data quality and task complexity on decision performance. Proceedings of the IRI—2005 IEEE International Conference on Information Reuse and Integration. IEEE. p. 149–54.

Juran JM, (ed.). 1951. Quality-control handbook. 1st ed. New York (NY): McGraw-Hilll.

Kawrykow A, Roumanis G, Kam A, Kwak D, Leung C, Wu C, Zarour E, Sarmenta L, Blanchette M, Waldispu hl J. 2012. Phylo: a citizen science approach for improving multiple sequence alignment. PLoS One 7:e31362.

Keil R, Salemme K, Forrest B, Neibauer J, Logsdon M. 2011. Differential presence of anthropogenic compounds dissolved in the marine waer of Puget Sounds, WA and Barkley Sound, BC. Mar Pollut Bull 62:2404–11.

Kelling S, Johnston A, Hochachka WM, Iliff M, Fink D, Gerbracht J, Lagoze C, La Sorte FA, Moore T, Wiggins A. 2015. Can observation skills of citizen scientists be estimated using species accumulation curves?. PLoS One 10:e0139600.

Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popovic Z, Baker D, Players F. 2011. Algorithm discovers b protein folding game players. Proc Natl Acad Sci U S A 108:18949–53.

Kim S, Robson C, Zimmerman T, Pierce J, Haber EM. 2011. Creek watch: pairing usefulness and usability for successful citizen science. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). New York (NY): ACM. p. 2125–34.

Kosmala M, Wiggins A, Swanson A, Simmons B. 2016. Assessing data quality in citizen science. Front Ecol Environ 14:551–60.

Kremen C, Ullman KS, Thorp RW. 2011. Evaluating the quality of citizen-scientist data on pollinator communities. Conserv Biol 25:607–17.

Lintott CJ, Schawinski K, Slosar A, Land K, Bamford S, Thomas D, Raddick MJ, Nichol RC, Szalay A, Andreescu D, et al. 2008. Galaxy zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Mon Not R Astron Soc 389:1179–89.

Marshall PJ, Verma A, More A, Davis CP, More S, Kapadia A, Parrish M, Snyder C, Wilcox J, Baeten E, et al. 2016. SPACE WARPS—I. Crowdsourcing the discovery of gravitational lenses. Mon Not R Astron Soc 455:1171–90.

Masters K, Oh EY, Cox J, Simmons B, Lintott C, Graham G, Greenhill A, Holmes K. 2016. Science learning via participation in online citizen science. J Sci Commun 15:A07.

Montecino-Latorre D, Eisenlord ME, Turner M, Yoshioka R, Harvell CD, Pattengill-Semmens CV, Nichols JD, Gaydos JK. 2016. Devastating transboundary impacts of sea star wasting disease on subtidal asteroids. PLoS One 11:e0163190.

Moyer-Horner L, Smith MM, Belt J. 2012. Citizen science and observer variability during American pika surveys. J Wildl Manage 76:1472–9.

Palmer JRB, Oltra A, Collantes F, Delgado JA, Lucientes J, Delacour S, Bengoa M, Eritja R, Bartumeus F. 2017. Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. Nat Commun 8:916–29.

Pandya R. 2012. A framework for engaging diverse communities in citizen science in the US. Front Ecol Environ 10:314–7.

Parrish JK, Litle K, Dolliver J, Hass T, Burgess H, Frost E, Wright C, Jones T. 2017. Defining the baseline and tracking change in seabird populations: the Coastal Observation and Seabird Survey Team (COASST). 2017. In: Cigliano JA, Ballard HL, editors. Citizen science for coastal and marine conservation. New York (NY): Routledge. p. 19–38.

Paul K, Quinn MS, Huijser MP, Graham J, Broberg L. 2014. An evaluation of a citizen science data collection program for recording wildlife observations along a highway. J Environ Manage 139:180–7.

Pillay R, Miller WAW, Hines JE, Joshi AA, Madhusudan MD. 2014. Accounting for false positives improves estimates of occupancy from key informant interviews. Divers Distrib 20:223–35.

Pocock MJO, Tweddle JC, Savage J, Robinson LD, Roy HE. 2017. The diversity and evolution of ecological and environmental citizen science. PLoS One 12:e0172579.

Powell J, Nash G, Bell P. 2013. GeoExposures: documenting temporary geological exposures in Great Britain through a citizen-science website. Proc Geol Assoc 124:638–47.

Raddick MJ, Bracey G, Gay PL, Lintott CJ, Murray P, Schawinski K, Szalay AS, Vandenbert J. 2010. Galaxy zoo: exploring the motivations of citizen science volunteers. Astron Educ Rev 9:010103.

Riesch H, Potter C. 2014. Citizen science as seen by scientists: methodological, epistemological and ethical dimensions. Public Underst Sci 23:107–20.

Rosemartin AH, Crimmins TM, Enquist CAF, Gerst KL, Kellermann JL, Posthumus EE, Denny EG, Guertin P, Marsh L, Weltzin JF. 2014. Organizing phenological data resources to inform natural resource conservation. Biol Conserv 173:90–7.

Rotman D, Preece J, Hammock J, Procita K, Hansen D, Parr C, Lewis D, Jacobs D. 2012. Dynamic changes in motivation in collaborative citizen-science projects. Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12). New York (NY): ACM. p. 217–226.

Seibert J, Etter S, Strobl B, Vis M, van Meerveld HJ. 2017. Can citizens observe what models need?—Evaluation of the potential value of crowd-based hydrological observations. 19th EGU General Assembly Conference Abstracts, Proceedings from the Conference held 23–28 April, 2017 in Vienna, Austria. p. 11620.

Stevens M, Vitos M, Altenbuchner J, Conquest G, Lewis J, Haklay M. 2014. Taking participatory citizen science to extremes. IEEE Pervasive Comput 13:20–9.

Sullivan BL, Aycrigg JL, Barry JH, Bonney RE, Bruns N, Cooper CB, Damoulas T, Dhondt AA, Dietterich T, Farnsworth A, et al. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. Biol Conserv 169:31–40.

Swanson A, Kosmala M, Lintott C, Packer C. 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. Conserv Biol 30:520–31.

Theobald EJ, Ettinger AK, Burgess HK, DeBey LB, Schmidt NR, Froehlich HE, Wagner C, HilleRisLambers J, Tewksbury J, Harsch MA, et al. 2015. Global change and local solutions: tapping the unrealized potential of citizen science for biodiversity research. Biol Conserv 181:236–44.

Thorson JT, Scheuerell MD, Semmens BX, PattengillSemmens CV. 2014. Demographic modeling of citizen science data informs habitat preferences and population dynamics of recovering fishes. Ecology 95:3251–8.

Ward DF. 2014. Understanding sampling and taxonomic biases recorded by citizen scientists. J Insect Conserv 18:753–6.

Wiggins A, Crowston K. 2011. From conservation to crowdsourcing: a typology of citizen science. Proceedings of the Forty-Fourth Hawaii International Conference on System Science (HICSS-44), Koloa, HI. IEEE.

Wiggins A, Bonney R, LeBuhn G, Parrish JK, Weltzin J. 2018. A science products inventory for citizen-science planning and evaluation. BioScience 68:436–44.

Wilderman CC, Barron A, Imgrund L. 2004. From the field: service provider's experience with two operational models for community science. Community-Based Collaboratives Research Consortium Journal. Spring 2004.

Wright DR, Underhill LG, Keene M, Knight AT. 2015. Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. Soc Nat Resour 28:1013–29.